

October 27, 2023

E-Filed

The Honorable Jon S. Tigar
United States District Court for the Northern District of California
San Francisco Courthouse
450 Golden Gate Avenue
San Francisco, California 94102

Re: *Doe I, et al. v. GitHub, Inc., et al.*; Lead Case No. 4:22-cv-06823-JST

Dear Judge Tigar:

Pursuant to the Court's September 15, 2023 Further Case Management Conference (ECF No. 155), Defendants GitHub, Inc. ("GitHub"); Microsoft Corporation ("Microsoft"); and OpenAI, Inc.; OpenAI, LP; OpenAI OpCo, LLC; OpenAI GP, LLC; OpenAI Startup Fund GP I, LLC; OpenAI Startup Fund I, LP; and OpenAI Startup Fund Management, LLC. ("OpenAI") (collectively, "Defendants") and Plaintiffs Does 1-5 hereby submit this joint letter brief regarding the parties' respective ESI Order proposals. The parties met and conferred in good faith and were unable to reach an agreement on an ESI Order.

Defendants' Position

Pursuant to Section G of the Standing Order for All Civil Cases Before District Judge Jon S. Tigar, Defendants used the "Model Stipulation & Order Re: Discovery of Electronically Stored Information for Patent Litigation" as the basis for their Proposed ESI Order. The model ESI order for patent cases is the appropriate starting point because of the technical subject matter of this case (and, indeed, the parties similarly used the model order for patent cases when drafting the protective order eventually entered as ECF Nos. 63 & 145).

The Declaration of Rose S. Lee ("Lee Declaration") explains each modification Defendants made to the model order. Attached as Exhibit 1 to the Lee Declaration is a clean copy of Defendants' Proposed ESI Order. Exhibit 2 to the Lee Declaration is a redline comparing Defendants' Proposed ESI Order with the model order.

Defendants attempted to modify the model order as little as possible. They made changes only when those changes were necessary because of the nature of the case or to accommodate requests made by Plaintiffs. For example, Defendants made changes to clarify how the multiple OpenAI Entities named as defendants will be treated for purposes of document collection and production, as well as to require the parties to meet and confer regarding the production of certain types of information for which the standard form of production would be inappropriate.

Plaintiffs, on the other hand, ask this Court to enter a 46-page proposed order that bears little resemblance to the model order, and goes far beyond what is reasonable or required by the Federal Rules of Civil Procedure. For example:

The Honorable Jon S. Tigar
October 27, 2023

Page 2

- Section III.A. of Plaintiffs' proposal requires that "[t]he Priority Document Custodians must include each Defendant's CEO." There is no justification at this stage of litigation for demanding that Defendants collect and produce documents from their most apex executives before Defendants produce documents from their most relevant custodians.
- Section III.B. of Plaintiffs' proposal adopts an entirely new protocol for handling discovery disputes (e.g., "If the Parties have not agreed whether to add the document custodian or source within 30 days of the requesting Party's request, then the matter may be brought to the Court in accordance with the procedures provided herein."). These procedures are inconsistent with this Court's standing orders on discovery motion practice, which requires that a party first obtain leave of court to submit a discovery motion, and otherwise meet and confer and submit a joint letter brief. (*See* Standing Order for All Civil Cases Before District Judge Jon S. Tigar, Section G.)
- Section V.B. of Plaintiffs' proposal provides that "[g]enerally, the Producing Party must produce in native format" documents including email and Word files. Production of all documents in native format would be extremely challenging in a world where ESI is readily altered. It is important to be able to ensure the contents of produced material in order to guarantee authenticity, particularly when parties litigating before this Court customarily stipulate to the authenticity of anything produced from a party's files and created by them. This would be impossible without the standard procedures of document imaging and Bates labeling. To address any legitimate concerns, Defendants have agreed to native format for those categories of ESI that are often unusable without it.
- Section IX of Plaintiffs' proposal is particularly burdensome and goes far beyond what is required by the Federal Rules of Civil Procedure. For example, Section IX.A. requires that a custodian's email be produced "regardless of whether such emails are in the custodian's actual email account." It is unclear how a reasonable search of all such emails could be conducted. Section IX.B. requires collection of each custodian's social media "messages, posts, direct messages, emails, video, and text messages." The scope of such collection requires Defendants to go far beyond the actions of custodians in their corporate capacities to collect and review all material generated in their personal capacities. This is not only highly burdensome but invades privacy rights. Meanwhile, Plaintiffs have not articulated any reason why personal information needs to be collected and produced, or how such information is relevant to any of the claims asserted. Sections IX.C., IX.D., and IX.G. require disclosure of burdensome cell phone data, including a list of cell phone numbers (including temporary or pay-by-month cell phones), *all necessary passwords or password information* for accessing the phone, a list of installed communications-related applications, logs of any calls and voicemails, all text messages, a list of contacts exported in Excel, and "billing statements for all phone number [*sic*] used by custodians or provide written authorizations for the release of such phone records from the applicable phone carriers." Finally, Section IX.K. states that "[a]n agreement to produce documents for a particular custodian shall include the

The Honorable Jon S. Tigar
October 27, 2023

Page 3

collection of that custodian's current and/or former administrative assistants," regardless of the relevancy of such assistants' documents.

The scope of collection and production proposed by Plaintiffs is staggering, unworkable, and inconsistent with the Court's model order. After the parties met and conferred, in an effort to reach agreement where possible, Defendants incorporated several of Plaintiffs' provisions, including (1) Plaintiffs' request that PowerPoint files be produced in native format; (2) Plaintiffs' request that ESI items be processed to preserve the date/time shown in the document as it was last saved rather than the date of collection or processing; and (3) Plaintiffs' provision regarding procedures for non-party subpoenas, including timely notifying other parties when non-party productions are received and providing copies of such productions to the other parties. The remainder of Plaintiffs' proposed order, however, is either already addressed in Defendants' proposed order or unreasonable.

Because Defendants' proposal is based on this District's model stipulated order in compliance with this Court's standing order, and Plaintiffs' 46-page proposal is overly burdensome and goes beyond the Federal Rules of Civil Procedure, this Court should enter Defendants' Proposed ESI Order.

Plaintiffs' Position

On September 15, 2023, this Court set a deadline of October 27, 2023 for the Parties to submit a stipulation regarding ESI. On that same day, Plaintiffs shared a comprehensive ESI protocol. Defendants first responded to Plaintiffs regarding ESI on October 18, 2023, by sending an entirely different proposed ESI protocol, only commenting it was based on the Northern District's model order. The parties met and conferred on October 24, 2023.

E-discovery has continued to evolve since the 2006 amendments to the Federal Rules first included the phrase "electronically stored information." *See* Fed. R. Civ. P. 26(a) advisory committee's note to 2006 amendment. While the Model Order is a helpful guide, it is meant to be a starting point. *See Hardin v. Mendocino Coast Dist. Hosp.*, No. 17CV05554JSTTSH, 2019 WL 4256383, at *2 (N.D. Cal. Sept. 9, 2019) ("[A]n ESI protocol has to be specific or it doesn't mean anything."). The Northern District's ESI Guidelines were last updated in 2015, nearly a decade ago. Best practices have evolved to keep pace with experience in complex cases such as this, together with the advancement of ESI discovery technology and procedures consistent with Rule 26.

This case is much more complex than non-class patent or trade secret matters—there will be discovery into sales and marketing, cutting-edge technology, class issues, damages, and other issues common in complex cases. Plaintiffs' Proposed ESI Protocol reflects counsels' experience in such matters and is uniquely tailored to the specific needs of this case, while remaining consistent with the Northern District's Guidelines. Indeed, similar ESI Protocols have been adopted by district courts across the nation (including in this District) managing complex cases with voluminous discovery records. *See, e.g., In re Xyrem (Sodium Oxybate) Antitrust Litig.*, Case No. 3:20-md-02966-RS (N.D. Cal.), ECF Nos. 80, 219; *In re Jan. 2021 Short Squeeze*

The Honorable Jon S. Tigar
October 27, 2023

Page 4

Trading Litig., Case No. 21-2989-MDL-ALTONAGA/Torres (S.D. Fla.), ECF No. 341; *In re Broiler Chicken Antitrust Litig.*, Case No. 1:16-cv-08637 (N.D. Ill.), ECF No. 459; *In re Generic Pharms. Pricing Antitrust Litig.*, Case No. 16-MD-2724 (E.D. Penn.), ECF Nos. 938, 1045; *In re Pork Antitrust Litig.*, Case No. 0:18-cv-01776-JRT-HB (D. Minn.), ECF No. 292; *In re Cattle Antitrust Litig.*, Case No. 0:19-cv-1222 (JRT/HB) (D. Minn.), ECF Nos. 93, 572.

Number of Custodians (Sections III.A–B): Defendants propose five custodians for each party. Plaintiffs initially requested 25 initial priority custodians but reduced that to 10. Plaintiffs have asked for organizational charts to assist in the identification of custodian, but none have been produced. Five custodians in total will be entirely inadequate for a case of this scale and type.

Listed ESI Categories (Section II.B): As used in the Rules, ESI is broadly defined. Fed R. Civ. Proc. 34(a)(1) advisory committee notes to 2006 amendment. Plaintiffs’ proposal lists specific types of ESI to ensure they are included. It does not require that all categories listed *will* be produced. Plaintiffs’ proposed Section XIV sets forth a reservation of rights memorializing this intent.

Specifications for Production Format (e.g., Native or Other) (Section V.B): Plaintiffs’ proposal seeks to specify categories of ESI for which native production would be consistent with Rule 26 and for which the preservation of metadata is especially important. This is consistent with the practice of courts in this district. *See, e.g., In re Facebook, Inc. Consumer Priv. User Profile Litig.*, 2021 WL 10282215, at *29 (N.D. Cal. Sep. 29, 2021); *In re Facebook PPC Advert. Litig.*, 2011 WL 1324516, at *2 (N.D. Cal. Apr. 6, 2011). Courts regularly order documents other than Excel spreadsheets and PowerPoint presentations produced in native format. *See, e.g., Cohen v. Apple Inc.*, 2020 WL 13582130, at *2 (N.D. Cal. Aug. 3, 2020).

Technology-Assisted Review (“TAR”) (Section IV.C): Plaintiff’s proposal includes a provision setting forth procedures for use of TAR, should a party wish to use it in lieu of human review. In that event, there should be a dialogue, considering the additional challenges associated with TAR review. The procedures Plaintiffs propose will allow the parties and, if necessary, the Court, to proceed efficiently and to evaluate whether use of TAR is warranted should a party wish to employ it. *Alivecor, Inc. v. Apple, Inc.*, No. 21-CV-03958-JSW, 2023 WL 2224431, at *1 (N.D. Cal. Feb. 23, 2023).

Procedures for Search-Term Negotiation (Section IV.E): Plaintiffs propose procedures for negotiating the use of search terms to identify responsive documents. Considering the complexity of the case and disputes that have occurred thus far, Plaintiffs do not expect the parties to negotiate search terms without disagreement. Plaintiffs merely wish to set forth ground rules for raising objections and resolving any disputes, which is important here given that to-date, only a smattering of documents have been produced by Defendants. *See Cohen*, 2020 WL 13582130, at *1.

Redaction Criteria Needed—No Relevancy Redactions (Section V.H): Plaintiffs include a provision setting forth criteria for redactions that prohibit redactions based only on relevancy. Such redactions are routinely rejected by courts in this district. *See, e.g., Shenwick v. Twitter*,

The Honorable Jon S. Tigar
October 27, 2023

Page 5

Inc., 2018 WL 833085, at *3 (N.D. Cal. Feb. 7, 2018); *In re High-Tech Emp. Antitrust Litig.*, 2013 WL 12230960, at *1 (N.D. Cal. Mar. 15, 2013).

CEOs as Priority Custodians (Section III.A): Plaintiffs’ proposal to include Defendants’ CEOs as document custodians reflects common practice in federal court e-discovery. Indeed, there is no rule in this Circuit that excludes CEOs from discovery. *See, e.g., Alta Devices, Inc. v. LG Elecs., Inc.*, 2019 WL 8757255, at *1 (N.D. Cal. Feb. 20, 2019); *Shenwick*, 2018 WL 833085, at *1; *see also, e.g., Page v. Bragg Cmties., LLC*, 2022 WL 17724407, at *2-3 (E.D.N.C. Dec. 15, 2022) (designating defendants’ founder and “past president” as a document custodian); *In re EpiPen Mktg., Sales Pracs & Antitrust Litig.*, 2018 WL 1440923, at *3 (D. Kan. Mar. 15, 2018);

Here, Defendants’ CEOs’ public statements confirm their personal involvement in AI-related matters and the fact they are percipient witnesses. For example, Microsoft’s CEO Satya Nadella testified mere weeks ago at the trial in *United States v. Google LLC*, No. 20-cv-03010 (D.D.C.) about the future of AI—stating in open-court, among other things, “I worry a lot, even in spite of my enthusiasm, that there is a new angle with AI.” OpenAI’s CEO Sam Altman has frequently discussed AI in interviews and public statements, including testimony before the U.S. Senate. *See* <https://www.judiciary.senate.gov/committee-activity/hearings/oversight-of-ai-rules-for-artificial-intelligence>. It is therefore likely these top executives possess discovery ESI, which can be objectively tested through the employment of search terms and hit reports.

Preservation Steps (Section VI.A): Plaintiffs include a provision requiring a responding party to describe its preservation efforts. This is in line with the Northern District Guidelines, which impose an obligation on the parties to “discuss preservation.” Guideline 2.01. The Sedona Conference also contemplates the parties discuss preservation. *See, e.g., The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 Sedona Conf. J. 1, 139 (2018).

Family Relationships: Identification and Inclusion (Section VII.A–B): Plaintiffs propose a provision requiring inclusion of non-responsive family members and other details regarding handling of family documents. Such a requirement is commonplace to contextualize documents and understand fully email and other communications. *See, e.g., In re Meta Pixel Healthcare Litig.*, 2023 WL 4361131, at *1 (N.D. Cal. June 2, 2023); *see also In re StubHub Refund Litig.*, 2023 WL 3092972, at *1-2 (N.D. Cal. Apr. 25, 2023).

Parameters for Custodial ESI Collection (IX): Plaintiffs propose a provision requiring production of certain sources of relevant ESI about the parties’ custodians. Without this default information for all custodians, the parties will be forced to litigate about what is relevant and proportional for each custodian separately.

The Honorable Jon S. Tigar
October 27, 2023

Page 6

Dated: October 27, 2023

By: /s/ Joseph R. Saveri
Joseph R. Saveri

Joseph R. Saveri (State Bar No. 130064)
Steven N. Williams (State Bar No. 175489)
Cadio Zirpoli (State Bar No. 179108)
Elissa A. Buchanan (State Bar No. 249996)
Travis Manfredi (State Bar No. 281779)
JOSEPH SAVERI LAW FIRM, LLP
601 California Street, Suite 1000
San Francisco, California 94108
Telephone: (415) 500-6800
Facsimile: (415) 395-9940
Email: jsaveri@saverilawfirm.com
swilliams@saverilawfirm.com
czirpoli@saverilawfirm.com
eabuchanan@saverilawfirm.com
tmanfredi@saverilawfirm.com

Counsel for Plaintiffs and the Proposed Class

Dated: October 27, 2023

By: /s/ Annette L. Hurst
Annette L. Hurst

Annette L. Hurst (State Bar No. 148738)
ORRICK, HERRINGTON & SUTCLIFFE LLP
The Orrick Building
405 Howard Street
San Francisco, CA 94105
Telephone: 415 773 5700
Facsimile: 415-773-5759
Email: ahurst@orrick.com

Counsel for GitHub, Inc. and Microsoft Corporation

Dated: October 27, 2023

By: /s/ Allyson Roz Bennett
Allyson Roz Bennett

Allyson Roz Bennett
MORRISON & FOERSTER LLP
707 Wilshire Boulevard, Suite 6000
Los Angeles, CA 90017-3543
(213) 892-5200
Fax: (213) 892-5454
Email: ABennett@mofo.com

The Honorable Jon S. Tigar
October 27, 2023

Page 7

Joseph C. Gratz (State Bar No. 240676)
MORRISON & FOERSTER LLP
425 Market Street
San Francisco, CA 94105
Telephone: 415-268-7000
Facsimile: 415-268-7522
Email: JGratz@mofo.com

*Counsel for OpenAI, Inc.; OpenAI, L.P.; OpenAI GP, L.L.C.;
OpenAI Startup Fund GP I, L.L.C.; OpenAI Startup Fund I,
L.P.; and OpenAI Startup Fund Management, LLC*

Attestation of Concurrence

I, Joseph R. Saveri, am the ECF User whose identification and password are being used to file the foregoing Joint Letter. Pursuant to Civil Local Rule 5-1(h)(3), I attest that concurrence in the filing of this document has been obtained.

Dated: October 27, 2023

JOSEPH SAVERI LAW FIRM, LLP

By: /s/ Joseph R. Saveri
Joseph R. Saveri